7-1-2013

Business Process Mining Success

Ronny Mans
Eindhoven University of Technology, Eindhoven, Noord-Brabant, Netherlands, r.s.mans@tue.nl

Hajo Reijers
Eindhoven University of Technology, Eindhoven, Noord Brabant, Netherlands, h.a.reijers@tue.nl

Hans Berends
Free University of Amsterdam, Amsterdam, Noord Holland, Netherlands, j.j.berends@vu.nl

Wasana Bandara
Queensland University of Technology, Brisbane, Queensland, Australia, w.bandara@qut.edu.au

Rogier Prince
Deloitte Consulting, Amstelveen, Noord Holland, Netherlands, rogprince@deloitte.nl

BUSINESS PROCESS MINING SUCCESS

Mans, Ronny, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB, Eindhoven, The Netherlands, r.s.mans@tue.nl[1]

Reijers, Hajo, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB, Eindhoven, The Netherlands, h.a.reijers@tue.nl

Berends, Hans, VU University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands, j.j.berends@vu.nl

Bandara, Wasana, Queensland University of Technology, Brisbane, Queensland 4001, Australia, w.bandara@qut.edu.au

Prince Rogier, Deloitte Consulting, Laan van Kronenburg 2, NL-1183 AS, Amstelveen, The Netherlands, rogprince@deloitte.nl

Abstract

Process mining has developed into a popular research discipline and nowadays its associated techniques are widely applied in practice. What is currently ill-understood is how the success of a process mining project can be measured and what the antecedent factors of process mining success are. We consider an improved, grounded understanding of these aspects of value to better manage the effectiveness and efficiency of process mining projects in practice. As such, we advance a model, tailored to the characteristics of process mining projects, which identifies and relates success factors and measures. We draw inspiration from the literature from related fields for the construction of a theoretical, a priori *model. That model has been validated and re-specified on the basis of a multiple case study, which involved four industrial process mining projects. The unique contribution of this paper is that it presents the first set of success factors and measures on the basis of an analysis of real process mining projects. The presented model can also serve as a basis for further extension and refinement using insights from additional analyses.*

Keywords: process mining, success factors, success measures, case study

[1] This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs.

1 Introduction

Process mining is a research discipline that is concerned with the discovery, monitoring, and improvement of operational processes by extracting knowledge from event logs as generated by information systems (Aalst, 2011a). The first publications on this subject appeared by the end of the 1990s (Agrawal et al, 1998) and since then many researchers have been working in this area. Process mining capabilities are nowadays offered by various academic tools (e.g. the ProM toolkit), as well as commercial systems (e.g. ReflectOne, Fujitsu Interstage, BusinesScape Enterprise Visualization Suite). It is fair to say that process mining has become a ‘hot topic’ in BPM research and spurs a huge interest from industry too (Aalst, 2011b).

In industrial practice, process mining is clearly not an end in itself. It is applied to discover the actual behavior of a business process, either in the absence of such an understanding or to check assumptions or expectations on that behavior (Aalst et al, 2007). Such insights are precious for organizations, since these can be used to improve that business process. At the same time, process mining is comparable with other business analysis activities in that it requires organizational investments. These include, for example, gaining authorization to access event logs, extracting the needed data from actual systems, preparing the event logs for process mining, configuring and using process mining techniques, and validating and interpreting process mining results (Aalst, 2011b).

To justify these organizational investments in practice, it is important to demonstrate the value and efficacy of a process mining project (cf. (DeLone et al, 2003)). An important step forward in this regard is the accomplishment of well-defined success outcome *measures*, which may be used to gauge the success of these investments. Once identified, the next step is to determine which of the characteristics of process mining projects can be identified as being connected to such measures. This may enable organizations to focus on the *success factors* that allow them to perform their process mining projects in an efficient and effective manner. Until now, only fragmented insights into process mining success measures and factors exist, and no systematic investigation has taken place into this topic. Our work is the first attempt to fill that gap.

This paper's contribution is a theoretically motivated and empirically validated model of success measures and factors for process mining projects. It is our belief that this model can raise awareness on the relevant and critical ingredients of process mining projects and, as such, increase their success rate in practice. One of the main sources for the proposed model is the work by Bandara et al. (Bandara et al, 2005;Bandara et al, 2006), who have looked into success measures and factors for *process modeling* projects. The motivation to build on this model is that the goals of process mining projects and process modeling projects often overlap, in the sense that they aim for the development of an understanding of a business process. As will be discussed in this paper, theoretical concepts from other domains, as well as practical insights, have been included in the final model to tune it towards the specific characteristics of process mining projects. In particular, our research has incorporated a multiple case study to validate the proposed model.

Against this backdrop, the structure of the paper is as follows. In Section 2 we will provide some theoretical background. Section 3 describes the research method we followed, including the case studies that were included. In Section 4, we will present our model for process mining success measures and factors, based on our findings. Finally, Section 5 includes a reflection on the limitations of our study and its implications.

2 Background

A basic requirement to apply process mining is the availability of an event log, which may originate from a wide variety of information systems that support business processes. For example, several projects have focused on the mining of data from ERP systems, such as a purchasing process for the sake of auditing and to test control objectives (Segers, 2007); a procurement and billing process within SAP R/3 (Ingvaldsen et al, 2008); and a fictive purchasing process in order to identify the potentials for doing process mining within SAP R/3 (Giesel, 2004). Other work has focused on the application of mining in the healthcare domain, such as: the discovery of a gynaecological oncology healthcare process (Mans et al, 2008); identifying the workflow of a field service engineer for diagnosing faulty X-ray machines (Rusu, 2010); and the flow of commands that are executed within X-ray machines (Günther et al, 2008). Furthermore, process mining approaches have been applied in multiple process improvement initiatives for the simulation of redesigned processes (Maruster et al, 2009); the identification of bottlenecks within IT service management processes (Driessen, 2006); and the discovery of software development processes (Hinojosa, 2008).

Despite the abundance of literature on process mining techniques (for an overview, see (Tiwari et al, 2008)) and the application of process mining techniques, to the best of our knowledge, no studies have been performed that identify and describe the success factors for process mining projects. Similarly, no studies are known to us which measure the overall success of process mining projects.

As mentioned in Section 1, we have an interest in both success measures and success factors. Success measures are criteria to evaluate the success of a project, whereas success factors are the influential forces that contribute directly or indirectly to the success of a project (cf. (Nemati et al, 2003)). Process mining has close links with a number of fields, hence it makes sense to reflect on the literature that is related to the successful execution of projects within these fields. We focus here on the general field of Information Systems (ISs), the field of data mining, and the field of process modeling (see Section 3).

An IS is a particular type of work system that uses information technology to capture, transmit, store, retrieve, manipulate, or display information, thereby supporting one or more other work systems (Alter, 1999). Regarding IS success, the DeLone and McLean success model (DeLone et al, 2003) is by far the most popular model (Sedera et al, 2002) and may be considered as the dominant basis of IS success measurement as many studies refer directly to this model (Urbach et al, 2009). The framework consists of six success dimensions that interact with one another: (1) 'information quality' which is a measure of the IS output; (2) 'system quality' which is a measure of the information processing system itself; (3) 'service quality' which is the provisioning of support for end user developers; (4) 'intention to use' which is the recipient consumption of the output of an IS; (5) 'user satisfaction' which is the recipient response to the use of an output of an IS; and (6) 'net benefits' which entail the impact of the IS on certain chosen stakeholders or aspects. Also, in literature, many works can be found regarding IS success factors (Bandara et al, 2005).

Process mining can be seen as a combination of data mining and process analysis (Aalst, 2011b). Data mining is defined in (Fayyad et al, 1996) as 'the application of specific algorithms for extracting patterns from data'. It is proposed in (Nemati et al, 2003) to use the square root (TSR) framework (Atkinson, 1999) for measuring data mining success, which integrates well-researched project success measures of the 'cost', 'time' and 'quality' triangle with IS success measures (e.g. maintainability), and organizational benefits (e.g. improved efficiency) as well as stakeholder community benefits (e.g. satisfied users). For organizational data mining (ODM) projects, key success factors are: (1) data quality, (2) data integration, (3) integration of technical components and expertise, (4) an ODM outsourcing strategy, (5) level of end-user ODM expertise, and (6) the Iron Triangle Elements. In (Feelders et al, 2000), the six factors are complemented by the statement that successful data mining projects require knowledge of data characteristics and the involvement of expertise in data mining, company data, and the subject area concerned.

Finally, there are close links between process mining and business process modelling (e.g. for process mining, the result is typically a process model). The latter can be defined as 'an approach for visually depicting how businesses conduct their operations; defining and depicting entities, activities, enablers, events, states and the relationships between them' (Bandara et al, 2005). In the process modeling field, the process modeling success model presented in (Bandara et al, 2005) is the only one known to us that rigorously deals with process modeling projects. It includes three success measures: (1) model quality, i.e. the extent to which all desirable properties of a model are fulfilled to satisfy the needs of the model users in an effective and efficient way; (2) process impacts, i.e. the overall effect of the initiative on the processes modeled; and (3) project efficiency, i.e. the ratio of obtained outcomes over invested resources. Additionally, several success factors are mentioned which influence the success of a process modeling project. These factors are divided into project specific factors (top management support, project management, and resource availability), modeling related factors (modeler expertise and modeling aids), and moderating factors (importance of the modeling project and complexity of the modeled process).

In summary, despite the abundance of literature on process mining algorithms and the wide application of process mining, no work on the successful execution of process mining projects exist. Clearly, in related fields, relevant insights exist.

3 Method

Given the lack of existing studies into process mining success, we set out to developing a model that captures success measures and success factors. In this section, we will explain the research method we followed, which covers the building of an *a priori* model and its subsequent re-specification and validation.

3.1 A Priori Model

The first part of our method deals with the so-called *a priori* identification of useful constructs for model building, cf.(Eisenhardt, 1989;Yin, 2003). For this purpose, a systematic literature review was conducted. In this review, literature of the IS, business process modeling and data mining fields, have been searched for success factors and success measures. Moreover, an analysis of the process mining literature led to the identification of 15 scholarly publications in which a real-life application of process mining was described. Applying grounded theory procedures to literature review (Wolfswinkel et al, 2011), these scholarly publications were analyzed to identify potential success measures and factors. Even though none of these publications specifically focus on process mining success, we expected that -- as a side-product of the application of process mining -- useful reflections on success measures and factors would be embedded in the discussions. To enable this identification, the set of 15 practice-oriented publications were imported in the qualitative data analysis tool NVivo (www.qsrinternational.com). Through ‘open coding’ those passages mentioning success factors and success measures were assigned to a node with a name closely related to the actual piece of coded text (Miles et al, 1984). To create a manageable set of measures and factors, similar nodes were regrouped in categories through a step called ‘axial-coding’ (Corbin et al, 2008). To prevent bias in the theoretical coding process, 5 of the 15 publications were analyzed by two coders, using a shared coding scheme, resulting in an inter-coder reliability of 85%. The coding was stopped when no additional concepts were found. As there are close links between process mining and business process modelling, the process modeling success framework as described in (Bandara et al, 2005;Bandara et al, 2006) was taken as a starting point. Subsequently, by refining it with success factors and success measures found in the literature, the identified concepts were used to establish a hypothesized *a priori* model of process mining success (see Figure 1- middle part). This conceptual model captures the hypothesized success factors and their relationships with the measures that can be used to evaluate process mining success. The rationale for including each construct is provided further in Sections 4.2 and 4.3, when the a priori constructs and re-specified constructs are discussed in detail (as based on case study data).

3.2 Multiple Case Study

The second part of our research method involved the execution of a multiple case study. Its purpose was to generate empirical insights that could be used to evaluate and possibly re-specify the *a priori* model. A case study design seemed appropriate given our interest in the natural, real-life application of process mining, and the scarcity of existing studies in the field (cf. (Benbasat et al, 1987)). A multiple case study design was chosen to circumvent critique on the lack of generalizability of single case studies (Bandara et al, 2005;Bandara et al, 2006). In the context of our study, a case was defined as *a single process mining project* that specifically aimed at the derivation of model of a business process (possibly among other goals, such as conformance checking). For the selection of cases, we applied theoretical sampling: Cases were chosen to fill theoretical categories. In other words, they should provide examples of ‘polar types’ to facilitate the development of theory with a broad application domain (Eisenhardt, 1989). In (Aalst, 2007), different types of ISs are classified in the context of process mining, which seemed useful to set up a theoretical categorization. One dimension, to which we shall refer as the ‘type of IT system’, distinguishes between *data centric* and *process centric* systems: IT systems respectively put the focus on the ordering of activities (e.g a workflow management system) versus sharing and exchange of data (e.g. an ERP system). The other dimension distinguishes between *structured* and *unstructured* systems, which differentiates between predefined and ad-hoc manners of dealing with work respectively. This latter dimension will be referred to as ‘process complexity’. Additional selection constraints were specified as follows: Stakeholders in the process mining project had to be available for interviews, and the project had to be conducted in the recent past (2007-2011). The case studies were all identified through our network of industrial contacts. This sampling process led to the inclusion of the four process mining projects we will discuss next:

- **Project A:** The project under consideration was carried out within ASML (http://www.asml.com), which is a large manufacturer of advanced technology systems for the semiconductor industry. The main aim of the project was to learn from highly detailed test log data how the actual process is executed to test wafer scanners before they are delivered to customers. The log data were derived from a single database, which contained the start and stop moment for calibrations, test actions, etc. This testing process is usually completely tailored to a specific instance and is further characterized by very low volumes (typically 50 systems within the same generation are produced).
- **Project B:** The process mining project within this case was carried out by the IT auditing department of a large German multinational company active in the energy, healthcare and manufacturing industries. The purpose of the project was to analyze the purchase-to-pay process, from purchase requisitions to outgoing payments. The event log was generated from the data that was captured by the company's SAP implementation. The process in question is characterized by high volumes of cases, which follow a reasonably predictable procedure.
- **Project C:** This process mining project took place within the Verbeeten Institute (http://www.verbeeten.nl), which is a specialized hospital with high expertise in radiotherapy and nuclear care. Under consideration was the institute's main process, which is the radiotherapeutic treatment of cancer. Event logs were generated from the workflow management system that supports this treatment process. It concerns a relatively linear process, which consists of different treatment stages that are consecutively planned per patient.
- **Project D:** This process mining project was conducted within the Dutch branch of T-Mobile (http://www.t-mobile.nl) to obtain more insight into one of their renewal processes. Specifically, the focus was on the activation of customer services for existing customers when they initiated a new iPhone subscription. A central database was in use to store information about customer contracts, orders, and order lines. While the operators are supported by front end systems to carry out a fairly standard procedure, a large number of exceptions were noted in this process; this was also what triggered the process mining project.

In Table 1, a mapping of these case studies to the theoretical categories is provided. Note that for all categories all values are covered, except that there is no case study that relates to the use of an *unstructured* yet *process centric* IT system. Considering that this is a rare combination in practice (Aalst, 2007)(p.15), this is not expected to have a negative impact on our goal to make a real-world assessment of process mining projects.

Project	Organization	Process	Type of IT system	Process complexity
A	ASML	Testing of wafer scanners	Unstructured & Data centric	High
B	Company B	Purchase-to-pay	Structured & Data Centric	Medium
C	Verbeeten	Radiotherapeutic treatment	Structured & Process Centric	Low
D	T-Mobile	Activation of customer services	Structured & Data Centric	High

Table 1: Classification of case studies.

For all case studies, two interviews were conducted with people with intimate knowledge of the process mining project. Our aim was to interview one person involved with the actual process mining, as well as to have an interview with either the project sponsor or project leader. Note that in one instance we were only able to interview people who were directly involved with the process mining analysis (project B). Each interview typically lasted for one hour. An interview protocol similar to that of (Bandara et al, 2005;Bandara et al, 2006) was followed, according to which open questions were used, first allowing the interviewee to comment on success factors and measures, and later bringing in the constructs of the *a priori* model for confirmation and comment. This was followed up with questions aimed at determining the importance of identified constructs. All interview transcripts were imported into Nvivo and stored in their respective case site folders. Next, a node structure was created in Nvivo to mirror the constructs of the *a priori* model. The coding of the interview transcripts then took place according to the following procedure:

- Coding only took place when relevant passages (mentioning something potentially related to success factors or measures) in the text were found.
- Passages of text that generally mentioned an existing construct in any way were coded under the relevant node of the *a priori* model. If new ideas or constructs were found that could not be easily placed under an existing construct, they were coded under a new node.
- When new ideas or constructs were found, previous transcripts were scanned again to look if all data relating to these new constructs was captured.
- General coded citations within each node were further investigated to distinguish between citations that stressed the *importance* or the explicit *unimportance* of constructs. Both positive and negative citations were coded as sub-nodes within the general construct.
- Generic citations were also screened for potential sub-constructs (things that people used to describe the construct), which contributed to further definition and target operationalization of these constructs.

3.3 Validation approach

The final part of our research method applied the case study findings to validate and, if required, to re-specify the *a priori* model. In line with (Bandara et al, 2005;Bandara et al, 2006), the case study data was analyzed to determine whether all important constructs were captured in the *a priori* model, as well as to identify constructs in the *a priori* model that are not critical or relevant. In addition, a reflection took place on the question whether any of the constructs in the model could be considered redundant.

When performing the cross-case analysis by comparing, for each construct across the different case sites, the number of general citations (merely mentioning the construct), important citations (mentioning the importance of the construct) and 'unimportance' citations (explicit mentions of the irrelevance of the construct) were compared. Coded data from the interviews yielded counts of the number of citations for each construct. The number of citations for each construct was summarized in a cross-case comparison matrix, which was created through a matrix coding query in Nvivo. To test for redundancy and relatedness of constructs, a different matrix intersection search was conducted with Nvivo. This served to identify passages that were coded under multiple constructs. In general, if a passage of text is coded under multiple constructs this may indicate a possible overlap or a moderating relationship. The Nvivo tool was used to check these specific instances in detail. The results of these procedures where taken together to enable the re-specification of the *a priori* model, based on the case study data (adapted after (Bandara et al, 2005;Bandara et al, 2006)). Following the same logic of (Bandara et al, 2005;Bandara et al, 2006), an *a priori* construct was to be deleted if it received only a few general citations, if it had strong citations for its irrelevance as a construct, and/or if the data coded under this construct was also coded under another.

4 Findings

In this section, we will present the finalized, re-specified process mining success model, which is visualized at the bottom of Figure 1. The overall research method is depicted in this figure: It can be seen how the *a priori* model (see middle part of Figure 1) was built first, based on the process modeling success model of Bandara (top) and the relevant literature (left and right).

This *a priori* model was then re-specified based on the findings of the case studies, resulting in the *re-specified* a priori process mining success model (bottom).

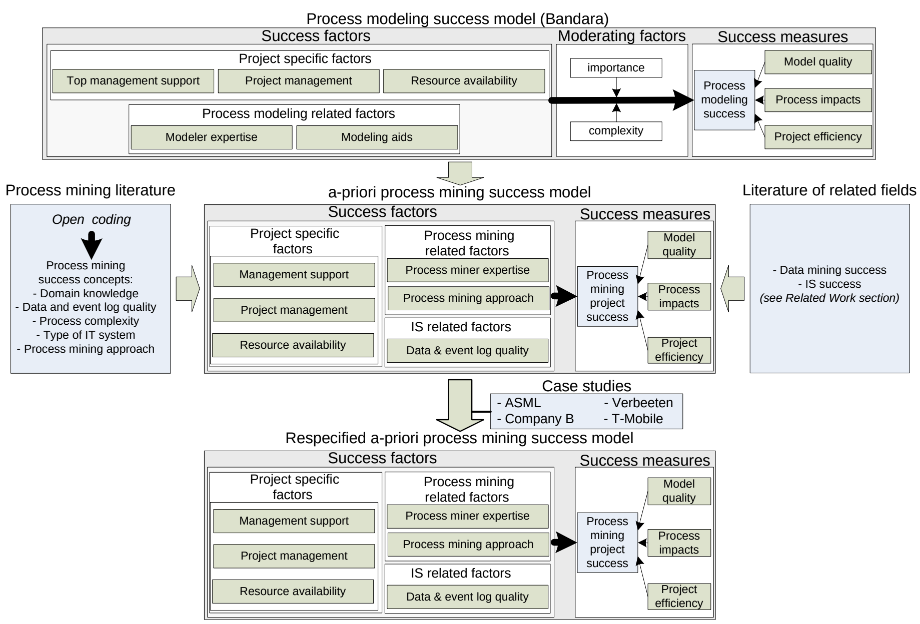

Figure 1: The re-specified a priori process mining success model.

In the remainder of this section, an overview of the *a priori* success model is provided. Next, we present on the actual re-specification of the *a priori* process mining success model on the basis of our interactions with the respondents from the case studies. Finally, a reflection on the findings is provided in Section 4.3.

4.1 Overview a priori model

The *a priori* process mining success model contains six success factors and three success measures. For the success factors, a distinction can be made between three different groups. The ‘Project specific factors’ group covers the three factors that hold for projects in general, whereas the ‘Process mining related factors’ group covers the two factors that are specific for process mining analysis. Finally, the ‘IS related factors’ group covers the factors that are specific for the data that originates from ISs, which heavily influence a process mining analysis.

The a priori process mining success model is different from Bandara’s success model in different respects. First, the ‘modeling aids’ success factor was changed into the ‘structured process mining approach’ success factor. The reason to focus on the mining approach is that for process mining relatively few tools exist, whereas for modeling a process many tools are available. To successfully apply process mining it seems more important to select the proper mining approach, as supported by a particular technique, than a particular tool – at least given the maturity of process mining tool support. Second, for the process mining success model the ‘IS related factors’ group and the ‘data and event log quality’ success factors have been newly defined, since process mining heavily relies on the data that originates from ISs. For process modeling, the data from ISs is much less relevant and for this reason did not appear in Bandara’s success model.

Also note that moderating variables of ‘importance’ and ‘complexity’ (as presented in the process modeling success model) were eliminated from the model even though they sometimes appeared in our data. The reason for this is that the number of cases available to support for the res-specification is too low to warrant a meaningful investigation of these contingency variables and their influence on model constructs.

4.2 Re-specification

In this section, we present a synopsis of our findings for the re-specification of the a-priori model. Full details of our findings can be found in our technical report (see (Mans et al, 2012)). In Table 2, a *definition* is given of each of the three success measures together with the *literature references* that were used for building the a-priori model (following the *a-priori* caption). Here, a distinction is made between references within the IS, business process modeling, and data mining fields (‘literature’ caption) and references within scholarly publications in which a real-life application of process mining was described (‘case studies’ caption). Furthermore, on the basis of our own case study findings we reflect on whether it would make sense to retain the success measures in the final, re-specified model. For the case study findings we will refer to the overall numbers of citations that particular constructs received over all case studies (shown in the *re-specification* column). A distinction is made between general citations (GC), citations that stress the importance of a construct (I), and citations that stress the non-importance of a construct (NI). Recall that general coded citations for a construct were further investigated to distinguish between citations that stressed the *importance* of the construct or its *lack of importance* being explicitly mentioned in this way. As a result, the sum of importance citations and non-importance citations are always lower or equal to the number of general citations.

A similar overview for the success factors is provided in Table 3. An additional piece of information that is added in this model is how the success factors are thought to influence the various success measures, according to the respondents (see ‘Influence’ caption).

A-priori	Respecification (own case study findings)			Decision
	GC	*I*	*NI*	

Model Quality: *The extent to which all desirable properties of a model created from process mining satisfy the needs of the model users*.				
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006))	18	12	3	In all positive citations, model quality was interpreted as the model's ability to provide relevant information and its conformance with user requirements. The negative citations took another perspective on quality, interpreting it as the technical accuracy of the model or its understandability. Therefore, the measure was kept but refined in accordance with these insights.
Case studies: process insights ((Günther et al, 2008;Hinojosa, 2008;Mans et al, 2008;Rusu, 2010))				
Process Impacts: *The overall effect of process mining on performance of a process.*				
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006)), data mining ((Nemati et al, 2003))	21	16	2	In two instances, it was indicated that insights to improve a process are not part of a process mining project. As the majority of the positive citations indicated that measuring process impacts are an important and relevant means for measuring the outcome of a process mining analysis, it was decided to keep the measure.
Case studies: no citations				
Project Efficiency: *The ratio of obtained outcomes over invested resources*.				
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006)), data mining ((Nemati et al, 2003))	10	8	1	Consistent, supporting citations were received across all projects. Therefore the measure was fully supported and retained in the model.
Case studies: no citations				

Table 2: Success Measures.

A-priori model	Respecification (own case study findings)			Influence	Decision
	GC	*I*	*NI*		
Management Support: *The involvement and participation of senior management, and their ongoing commitment and willingness to devote necessary resources and time of senior managers to oversee the process mining efforts*.					
Literature: process modeling projects ((Bandara et al, 2005;Bandara et al, 2006))	15	12	1	process impacts: increased uptake of the outcomes of the process mining analysis.	The factor was consistently cited across interviewees. One negative citation was received, but here the management support was provided due to an explicit audit mandate which means that all required resources need to be provided upon request. We decided to keep the factor in the model.
Case studies: no citations					
Project Management: *The management of activities and resources throughout all phases of the process mining project, to obtain the defined project outcomes*.					
Literature: process modeling ((Bandara et al, 2005)), data mining ((Nemati et al, 2003))	23	10	0	project efficiency: project completed in shorter time period.	Of the 23 citations, there were 10 importance citations and no non-importance citations. So, the factor was kept in the model.
Case studies: no citations					
Resource Availability: *The degree of information available from the project stakeholders during the entire process mining analysis*.					
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006)), data mining ((Feelders et al, 2000))	19	16	0	project efficiency: quicker performance of process mining analysis.	Of the 19 citations, there were 16 importance citations and no non-importance citations. The factor was kept in the model.

Case studies: (Aalst, 2005;Giesel, 2004;Ingvaldsen et al, 2008;Maruster et al, 2009;Segers, 2007)					
Process Miner Expertise: *The experiences of the person conducting the mining, in terms of event log construction, doing process mining analysis and knowledge of the business processes being mined*.					
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006))	31	26	0	model quality: less mistakes during analysis. project efficiency: project completed in shorter time period.	Of the 31 citations, there were 26 importance citations and no non-importance citations. So, the factor is fully supported and kept in the model.
Case studies: no citations					
Structured Process Mining Approach: *The extent to which a process miner uses a structured approach during the entire process mining analysis*.					
Literature: process modeling ((Bandara et al, 2005;Bandara et al, 2006)), data mining ((Fayyad et al, 1996))	12	10	0	model quality: increased usefulness and trustworthiness of the results.	Of the 12 citations, there were 10 importance citations and no non-importance citations. Again, the factor was decided to be kept in the model.
Case studies: no citations					
Data and Event Log Quality: *The characteristics of the raw data and subsequently constructed event logs*.					
Literature: data mining ((Nemati et al, 2003))	16	11	1	model quality: increased trustworthiness of process mining analysis outcomes.	The factor was consistently cited across interviewees and projects. Only for Company B both a positive and negative citation was received. The factor was kept in the model.
Case studies: (Driessen, 2006;Günther et al, 2008;Mans et al, 2008;Rusu, 2010)					

Table 3: Success Factors.

4.3 Reflection

By inspecting the presented tables, the reader can establish that all success measures and success factors in the *a priori* model received support from the respondents and were kept in the re-specified, final model. However, it is important to emphasize the following insights with respect to the success measures (see Table 2):

1. For the ‘model quality’ success measure, quality should be interpreted as the *information relevance* of the model and its *conformance with user requirements*. From the interaction with our respondents, we were able to conclude that model quality could also be interpreted as the *readability* of the model or the *technical fit* of the model with the algorithm’s purpose. These latter aspects were considered not to be important at all.
2. ‘Process impacts’ can be considered as an important success measure, but we learned that a distinction can be made between two kinds of projects. For one kind of project, process improvements *are* part of the project whereas for the other kind of project this is not. For the latter kind of projects we recommend that some effort is put into estimating the impacts of improvements which are derived from the obtaining process mining insights. This will help in showing the effects of possible improvements but it will also aid in appreciating the efforts and costs that are put in the process mining analyses.

For the success factors, it can be concluded from Table 3 that all success factors were retained in the re-specified model as well, and exactly in the way they were specified in the a priori model. It is noteworthy that the expertise of the process miner was the most cited success factor overall (31 general citations). The interviewees additionally indicated that during the entire process mining analysis, many mistakes can be made. The risk of making such mistakes is much reduced when an experienced process miner performs the analysis. Altogether, this suggests that the 'process miner expertise' success factor moderates the 'model quality' and 'project efficiency' success measures.

Finally, our analysis also made clear that each success factor has its influence on at least one success measure. In particular, among the interviewees, strong support has been found for the influence of the 'management support' success factor on the 'process impacts' success measure.

Also, for the 'structured process mining approach' and 'data and event log quality' success factors, it was found that they were unique for process mining as during the analysis a specific approach needs to be followed. For example, for the 'structured process mining approach' factor, it was found that for applying process mining many different types of algorithms are available, which underlines the importance of selecting the right ones during the process mining analysis endeavor.

For potential extension of the *a priori* model, the interviewees were also asked to propose additional, relevant success factors and successful measures. To this end, also the data transcripts were searched for new ideas and constructs. In total, three candidates for new success factors and two new candidates for success measures were explicitly mentioned , specifically 'personal skills', 'data privacy & authorization' and 'personal commitment' as success factors, and 'recognizability of results' and 'translating results to customer experience' as success measures. None of these were found more than once. Therefore, we decided to not go back to the interviewees and as a result none of the candidates were included in the final model.

5 Conclusions

Process mining is growing in popularity. It is used to discover, monitor, and improve real processes by extracting knowledge from event logs available in information systems. Although there is a plethora of publications on algorithms and the application of process mining, there is a dearth of knowledge and information about process mining success. In particular, to date there have been no studies that have investigated what the antecedent factors of process mining success are, or how to measure the success of a process mining effort. Addressing this gap has been the goal driving this study.

The study provided a validated re-specified a priori process mining success model consisting of three success measures and 5 success factors. Our findings are new in that they adapt Bandara's original modelling success model with success factors specific for process-mining, particularly 'structured process mining approach' and 'data and event log quality'. The re-specification phase strongly confirmed the validity of the success measures and factors in the a priori model, even though it implied (a) that 'model quality' must be interpreted with respect to usefulness rather than technical model quality and (b) that in auditing contexts 'process impacts' in terms of process performance improvement may not be that relevant – detection of a deviation may be sufficient. An additional insight that our study brings is the importance of 'process miner expertise' as a success factor for process mining projects. This is the factor that we expect to strongly moderate the 'model quality' and 'project efficiency' success factors, even though this needs to be confirmed in follow-up studies.

The study has several limitations. The a priori model was mainly derived from theory and input from analogous literature on domains different (though similar) to process mining. While this was addressed by a detailed model validation, there were inherent limitations in the case study phase as well. The results presented here were limited to four cases, where the analysis was based on the interviews of selected stakeholders (i.e. the process miner and project sponsor). The study was also prone to the more general limitations of case study research, such as case selection bias.

Despite these limitations, we believe that our results are of value. The study shows the commonalities and differences between process modeling and process mining projects. With respect to the commonalities, it seems useful to investigate how the plethora of publications on how to perform process modeling in practice can be exploited for improving the success of process mining projects (e.g. (Rosemann, 2006;Sharp et al, 2009)). Also, since our study is the first one forming a theory of process mining success, it may serve as a basis for future studies in which our tentative model can be extended and refined, and finally how to operationalize the constructs in the model.

References

Aalst, W. M. P. v. d. (2005). Business Alignment: Using Process Mining as a Tool for Delta Analysis and Conformance Testing. Requirements Engineering, 198-211.

Aalst, W. M. P. v. d. (2007). Exploring the CSCW Spectrum using Process Mining. Advanced Engineering Informatics, 191-199.

Aalst, W. M. P. v. d. (2011a). Process Mining Manifesto. Lecture Notes in Business Information Processing, 99, 169-194.

Aalst, W. M. P. v. d. (2011b). Process Mining: Discovery, Conformance and Enhancement of Business Processes. Berlin, Springer-Verlag.

Aalst, W. M. P. v. d., Reijers, H. A., Weijters, A. J. M. M., Dongen, B. F. v., Alves de Medeiros, A. K., Song, M. S., and Verbeek, H. M. W. (2007). Business Process Mining: An Industrial Application. Information Systems, 32 (5), 713-732.

Agrawal, R., Gunopoulos, D., and Leymann, F. (1998). Mining Process Models from Workflow Logs. EDBT 98.

Alter, S. (1999). Information Systems: A Management Perspective. Reading, Massachusetts, Addison Wesley.

Atkinson, R. (1999). Project Management: Cost, Time, Quality, Two Best Guesses and a Phenomenon, It's Time to Accept Another Success Critera. International Journal of Project Management, 17 (6), 337-342.

Bandara, W., Gable, G. G., and Rosemann, M. (2005). Factors and Measures of Business Modelling: Model Building Through a Multiple Case Study. European Journal of Information Systems, 347-360.

Bandara, W., Gable, G. G., Rosemann, M., Straub, D. W., and Klein, S. (2006). Business Process Modelling Success: An Empirically Tested Measurement Model. ICIS 2006.

Benbasat, I., Goldstein, D. K., and Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. MIS Quarterly, 11(3), 369-386.

Corbin, J. M. and Strauss, A. L. (2008). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, USA, Sage Publications.

DeLone, W. and McLean, E. R. (2003). The DeLone and McLean Model of IS Success. Journal of Management Information Systems, 19 (4), 9-30.

Driessen, S. (2006). Finding Process Bottlenecks from Logged Process Data using Data Mining. Master's Thesis (http://bit.ly/T556J6). Eindhoven, Eindhoven University of Technology.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. Academy of Management Review, 14 (4), 532-550.

Fayyad, U. M. and Piatetsky-Shapiro, G. (1996). From Data Mining to Knowledge Discovery in Databases. AI Magazine, 17 (3), 37-54.

Feelders, A., Daniels, H., and Holsheimer, M. (2000). Methodological and Practical Aspects of Data Mining. Information & Management, 37 (5), 271-281.

Giesel, M. v. (2004). Process Mining in SAP R/3. Master's Thesis (http://bit.ly/QgHW3h). Eindhoven, Eindhoven University of Technology.

Günther, C., Rozinat, A., and Aalst, W. M. P. v. d. (2008). Monitoring Deployed Application Usage with Process Miniing. BPM Center Report BPM-08-11.

Hinojosa, E. U. (2008). Process Mining Applied to the Change Control Board Process. Master's Thesis (http://bit.ly/StEEor). Eindhoven, Eindhoven University of Technology.

Ingvaldsen, J. and Gulla, J. (2008). Preprocessing Support for Large Scale Process Mining of SAP Transactions. Proceedings of BPM 2008.

Mans, R. S., Reijers, H. A., Berends, J. J., Bandara, W., and Prince, R. (2012). Business Process Mining Success - Extended Version. BPM Center Report BPM-12-17.

Mans, R. S., Schonenberg, M. H., Song, M. S., Aalst, W. M. P. v. d., and Bakker, P. J. M. (2008). Application of Process Mining in Healthcare - A Case Study in a Dutch Hospital. BIOSTEC 2008.

Maruster, L. and Beest, N. R. T. P. v. (2009). Redesigning Business Processes: A Methodology Based on Simulation and Process Mining Techniques. Knowledge and Information Systems, 21 (3), 267-297.

Miles, M. B. and Huberman, A. M. (1984). Qualitative Data Analysis: A Sourcebook of New Methods. Thousand Oaks, Massachusetts, USA, Sage publications.

Nemati, H. R. and Barko, C. D. (2003). Key Factors for Achieving Organizational Data-Mining Success. Industrial Management & Data Systems, 103 (4), 282-292.

Rosemann, M. (2006). Potential Pitfalls of Process Modeling: Part A. Business Process Management Journal, 12 (2), 249-254.

Rusu, S. (2010). Discovery and Analysis of Field Engineer Process Using Process Mining. Master's Thesis (http://bit.ly/WbnrWY). Eindhoven, Eindhoven University of Technology.

Sedera, W., Rosemann, M., and Gable, G. G. (2002). Measuring Process Modeling Success. ECIS 2002.

Segers, I. E. A. (2007). Investigating the Application of Process Mining for Auditing Purposes. Master's Thesis (http://bit.ly/TGUw6h). Eindhoven, Eindhoven University of Technology.

Sharp, A. and McDermott, P. (2009). Workflow Modeling: Tools for Process Improvement and Applications Development. Artech House Publishers.

Tiwari, A. and Turner, C. J. (2008). A Review of Business Process Mining: State-of-the-Art and Future Trends. Business Process Management Journal, 14(1), 5-22.

Urbach, N., Smolnik, S., and Riempp, G. (2009). The State of Research on Information Systems Success - A Review of Existing Multidimensional Approaches. Business & Information Systems Engineering, 1 (4), 315-325.

Wolfswinkel, J., Furtmüller, E., and Wilderom, C. (2011). Using Grounded Theory as a Method for Rigorously Reviewing Literature. European Journal of Information Systems, 1-11.

Yin, R. K. (2003). Case Study Research: Design and Methods. Applied Social Research Methods Series, 5. Thousand Oaks, Massachusetts, USA, Sage publications.